

Determining heterogeneous behavior for theater attendance

Victoria Ateca-Amestoy

Received: 2 March 2007 / Accepted: 28 April 2008 / Published online: 16 May 2008
© Springer Science+Business Media, LLC. 2008

Abstract In this paper we propose and estimate a model of theater participation using the data contained in the 2002 *Survey of Public Participation in the Arts* from the USA, a dataset widely used to study the determinants of cultural participation. Our contribution relies on the use of an estimation technique that respects the count data nature of the attendance variable (number of theater performances that an individual attended) and allows for heterogeneous behavior. By using a Zero Inflated Negative Binomial Model, we can characterize two distinct behaviors for the observable attendance: a group of never-goers (who never participate) and a subpopulation that has a positive probability of attending. For this latter group, we can estimate the effect of certain personal variables on the probability of highest frequency. The results suggest that the proposed model is appropriate for estimating cultural participation.

Keywords Cultural participation · Theater attendance · Cultural capital · Zero Inflated Negative Binomial Model

1 Introduction

Why is it that so few people attend theater performances? In 2002, 12.3% of the adult population of the United States attended a live theater performance, as

V. Ateca-Amestoy
Universidad del País Vasco and Instituto de Estudios Sociales Avanzados (IESA-CSIC),
Cordoba, Spain

V. Ateca-Amestoy (✉)
Department of Economic Analysis II, University of the Basque Country, Avda. Lehendakari
Aguirre, 83, 48015 Bilbao, Spain
e-mail: victoriamaria.ateca@ehu.es

recorded in the *Survey of Public Participation in the Arts* for that year.¹ Looking back in time, we can observe that participation rates remained quite stable for the period 1992–2002 (13.5% to 12.3% of the adult population) and that the total number of attendances decreased slightly for non-musical plays (by 2.49%, leading to a total of 58.7 million attendances in 2002). This did not happen with attendance of other performing arts or cultural activities—the so called “benchmark disciplines”—for which attendances increased (by 6.44% for musical plays, 16.43% for museums and art galleries, 20.49% for jazz, 20.73% for classical music, and up to 27.88% for opera performances). Only ballet attendance experienced a bigger decline over that period (down by 8.78%). This decline in audience for theater has continued on and professional theater organizations report an overall decline in audiences of about 8% for the period 2002–2006, despite increases in the number of performances offered over those years (by nearly 2%).²

When observing different participation profiles for different types of performing arts, one could be tempted to assign a special status to each cultural good, and feel that traditional consumer models may fail to accommodate such behavior; it could be that tastes differ greatly from one agent to another or that the changes in tastes in the course of a lifetime may imply an exception to diminishing marginal utility. However, assuming that tastes are stable and do not differ too much (Becker 1965; Stigler and Becker 1977), we can conclude that what drives a different observed behavior can be explained in terms of differences in constraints.

When determining the influence of different variables on human behavior regarding cultural goods, we should consider the gap between original preferences and the preferences revealed by the consumer's choice. We can present the argument following an example by Frey: an agent could have a great love for opera but reveal a low preference if he/she chooses to attend few or no live performances (Frey 2003). This divergence could be due to binding budget constraints, time constraints, social constraints, physical constraints, or low cultural capital endowments.

In this paper, we focus on theater participation in the United States and estimate a model that incorporates personal variables related with cultural capital to capture the influence of the availability of personal resources on attendance. We use a technique that accounts for unobserved heterogeneity and deals appropriately with the “excess zeros” observed in the dependent variable. This estimation method, to our knowledge, has not been used before to explain cultural participation.³ Taking advantage of the data contained in the 2002 *Survey of Public Participation in the Arts* (SPPA2002) for the United States, we use count data regression models to estimate the correlates of cultural participation by means of theater attendance. We respect the count data nature of a measure of attendance recorded in the survey and

¹ Survey data derived from the 1992 and the 2002 releases of the *Survey of Public Participation in the Arts*, as recorded in NEA Research Division Report #45 (NEA 2004).

² Data provided in the *Theater Facts 2006: A Report on Practices and Performance in the American Not-for-profit Theater Based on the Annual TCG Fiscal Survey*.

³ As we will discuss in the section that presents the estimation method, the same cannot be said of other areas such as the demand for health services or recreational services.

estimate a Zero Inflated Negative Binomial Model that allows us to jointly determine a participation function for two distinct groups of people.

This allows us to assume that there could be a given fraction of people whose optimal choice is never go to this kind of cultural activity, and thus assume that they reach an unconstrained maximum by never attending a theater performance. Another fraction of the total population has a positive probability of attending. The fact that a given agent did not attend during the period (the year in which the Survey from which we derive the data was run), could be due to the fact that he/she could not reach anything other than a corner solution if one of the constraints that determine his/her feasible choices is binding. Thus, the data generating process could be a mixture of these two heterogeneous processes. What could we say about the behavior of an individual from the first group? We could infer that there are some variables that affect his/her zero taste for this particular way of satisfying his/her cultural or leisure needs. Regarding the second group of people, we investigate a function that explains the frequency of attendance as a function of the same set of variables.

Our paper relies on a line of discussion that highlights the particular significance of cultural capital variables in explaining the differences in observed behavior; it is not only binding financial and time constraints that lead to corner solutions, but also scarcity of this type of personal resource. Cultural capital stock is a personal resource accumulated by past consumption of cultural goods (past exposure to cultural goods), so it makes individuals more productive in the fulfillment of their cultural needs. Closely related papers from theoretical and empirical literature explain the influence of past consumption capital on present demand, and the relevance of general human capital and specific human cultural capital in determining the profile of people who participate more in the arts (Stigler and Becker 1977; Gray 2003; Smith unpublished; Borgonovi 2004).

When considering the specific cultural good that we are analyzing—theater performances—we must take into account that it is not a material cultural good that can be integrated into a physical stock of capital. It seems to us that this is an especially interesting feature for testing any theoretical model of rational addiction and cultural capital formation, since this good is integrated into the consumption capital stock simply by means of a memory appreciation process. Theater performances are a good that consumers can purchase in the market at a given price. The number of tickets purchased in a given period of time depends on ticket prices, income, tastes, technology, and other restrictions (time constraints, the transportation cost associated with attending live shows, etc.). Unfortunately, as indicated by Borgonovi (2004), who enriches her analysis by introducing price related variables derived from another statistical source, the fact that ticket prices are not available forces the researcher to estimate participation functions in an attempt to determine the correlates of participation.

The next section comments on some of the approaches used and studies conducted on participation in the arts, with special emphasis on those that have used the information contained in SPPA2002. In Sect. 3, we describe the survey and the variables selected based on the behavioral model. Section 4 discusses testable hypotheses derived from (i) the cultural capital theoretical approach to the demand

for cultural goods, and (ii) a bundle of regularities reported by empirical studies. Section 5 proposes the estimation strategy and selects the most suitable estimation method from a bundle of count data regression models. As well as presenting the chosen estimation method Zero Inflated Negative Binomial Regression (ZINB), we discuss our reasons for using it (the model selection process is explained in more detail in the final appendix). Results are discussed in Sect. 6. Finally, Sect. 7 concludes.

2 Review of the literature

Other theoretical and empirical papers have studied the determinants of participation in theater performances; some of them have even used data derived from different releases of the SPPA. In this section, we review the concept of cultural participation and discuss different methodological approaches to estimating participation equations.

The literature on participation defines the concept as (i) “frequentation” (attendance at live performances; which is the relevant dimension for our analysis); (ii) the consumption of cultural goods through the media; and (iii) the practice of artistic activities. Participation is usually defined by the different ways in which it takes place.

When talking about the different theoretical approaches to participation in the arts, McCarthy et al. (2001a) point out the relevant contribution of cultural economics to the characterization of the determinants of individual behavior. Traditionally, Economics has provided decision-making models based on rational agents maximizing utility subject to their constraints. The simplest models consider no specific characteristics for cultural goods. Stigler and Becker (1977) offer a reformulation of these models with the introduction of personal cultural capital developed by the aggregation of past consumption levels. Under the presence of this production factor, individuals are more productive agents when they fulfill their cultural appreciation needs. Cultural capital and the variables that influence its law of motion play a key role in these models. Other relevant contributions came from institutionalism. For instance, Veblen considers non-functional demand represented by conspicuous consumption (Veblen 1965). For some authors (Bourdieu 1988), one of the main motivations for participation in high culture activities is the sense of belonging to a prosperous social class. Other approaches characterize the correlation between empirical variables and participation. For instance, interdisciplinary research on leisure discusses motivation and participation, the quantity and nature of leisure time and the social dimension of participation.

The empirical literature on participation in the arts describes consumption patterns in terms of three different dimensions: participation level, the characteristics of participants, and the determinants of participation. Though most empirical studies focus on who participates, rather than why they participate, the following factors are used to explain changes in participation patterns: socio-demographic changes, changes in tastes, changes in practical considerations (as supply-side

changes), leisure time availability, income levels, and changes in the cultural capital stock. Smith (unpublished) separates contributions on what he calls the “economic-demographic approximation” and on the “taste phenomenon approximation”. Contributions in the first group are based on the influence of prices, substitutes, income, and attributes, and a traditional economic approach is followed (Gapinski 1986; Moore 1966; DiMaggio and Ostrower 1990; Gray 1995). The second group highlights the significance of personal cultural capital variables, in the sense of “rational addiction” à la Stigler and Becker, or “learning-by-consuming” (as per Lévy-Garboua and Montmarquette 1996; Abbé-Decarroux and Grin 1992; Morrison and West 1986). Most of these studies treat participation as a dichotomous phenomenon, and estimate binary decision models. However, some authors have tried to explain the determinants, and the degree, of participation. One of the most comprehensive reviews of the empirical literature is the work by Seaman (2005). This author discusses in detail unsolved questions regarding participation, presents critically methods and results of the literature, and compares them.

Other authors, mostly sociologists, have used the information contained in SPPA2002 to study the demand for the arts. DiMaggio and Mukhtar (2004) study the factors that induced a change in consumption patterns for the period 1982–2002. Borgonovi (2004) estimates models to determine both participation and count for participation. However, these studies differ in their purposes and in the methods used to characterize demand for the arts in the USA. DiMaggio and Mukhtar focus on the change in participation rates by studying social groups under three different criteria: age, sex and education level. Borgonovi estimates a logit model for participation and an ordered logit model for attendance (count data are exogenously grouped into different categories). This last analysis shares some characteristics with the model that we propose, mainly in the selection of the relevant set of regressors.

The model to be estimated is by no means innocent: there are important behavioral assumptions that may rely on the suitability of our estimation method.⁴

The methodology employed by Fernández-Blanco et al. (2004) to estimate demand for cinema in Spain has the advantage of using latent class analysis. These authors estimate a finite mixture model, so different behavior for different subsamples can be estimated. If we believe that we cannot distinguish between never-goers and attendees, but can distinguish between groups with high average participation and low average participation, then this kind of estimation provides a good approach. Borgonovi (2004) justifies the estimation of a binary model for participation and an ordered model for frequency of participation on the grounds of the following assumption: there may be major differences characterizing the behavior of occasional attendees and frequent participants in the live arts. For this author, this is a call to study the 2 groups separately.

To our knowledge, no author has estimated participation equations by means of a count data model that deals with heterogeneity behind the no participation outcome.

⁴ Good examples of empirical exercises that explore the best way to estimate heterogeneity in behavior when using physician services are Jiménez-Martín et al. (2002) and Deb and Trivedi (2002). Methodological questions are covered in depth in Cameron and Trivedi (1998).

We believe that our proposed empirical model (which implies a latent classification of behavior and a joint estimation of both types of behavior) overcomes some of the problems with previous research.

3 Data and hypotheses

3.1 Description of the 2002 *Survey of Public Participation in the Arts*

This dataset compiles information on participation in the arts by US citizens between August 1st 2001 and August 1st 2002. It is the fifth study in a series conducted by the Bureau of the Census for the National Endowment for the Arts since 1982 and was conducted as a supplement to the *Current Population Survey*, a national survey on labor conditions (U.S. Bureau of the Census 2003). The dataset contains information on attendance at the following artistic activities: jazz, classical music, opera, musicals, theatre plays, ballet, dance, art museums, art crafts, and visits to historical parks and monuments. It also covers other types of cultural practice such as the consumption of cultural goods through the media, reading, active practice of artistic activities, education and early exposure to the arts, preferences for different types of music and school socialization. 17,135 questionnaires were completed for a representative sample of households in the USA. The sample was drawn up following a multi-stage strata design with clusters, based on information from the Bureau of the Census. In each of the selected households, all individuals over 18 were interviewed. A weighting variable makes the sample representative for the adult civilian population by age, sex, and ethnic origin.

3.2 Empirical specification and variables

We want to estimate a frequentation function using the count nature of the variable that allows us to validate empirically the results derived from analytical models and reported regularities. From analytical models, we know how the optimal demand depends on the socio-economic variables under a taste formation framework (Lévy-Garboua and Montmarquette 1996; Gray 2003; Smith unpublished; Ateca-Amestoy 2007). Cultural participation is determined by the availability of certain resources used to satisfy cultural needs (namely time, money, social capital, and cultural capital) as well as by personal and household characteristics.

Let us assume that the number of attendances (y_i) depends on the following variables.

$$y_i = f(\mathbf{x}_i) = f(S_i, P_i, O_i, Sc_i, H_i, Se_i, De_i, E_i) \quad (1)$$

where S_i represents variables related to the stock of available cultural capital; P_i captures preferences (in the form of a measure for constrained or unconstrained choice); O_i includes participation other than by attending live performances; Sc_i is the variable for social participation; H_i , Se_i , and De_i provide information about socio-economic and demographic characteristics of the individual and his/her household; and finally E_i stands for the geographical variable that allows us to

Table 1 Variables that enter into the analysis

Variables	Definition	Hypothesis
y_i	Number of theater performances over the last year	
S_i	Vector of cultural capital variables	
$edum_i$	Mother formal education level	+
$edup_i$	Father formal education level	+
edu_i	Own formal education level	+
$clas_i$	Theater classes (differentiated by starting age)	+
act_i	Acts	+
P_i	Vector of preferences	
$gomore_i$	Would go more if no restrictions	+
O_i	Participation in theater by means of other cultural goods	
$mediaact_i$	No live theater/active (video, DVD)	+/-
$mediapas_i$	No live theater/passive (TV, radio)	+/-
$readt_i$	Reads theater	+/-
Sc_i	Vector of social capital	
$familygo_i$	Participation of any other member of the family	+
H_i	Vector of household composition variables	
$hhold_i$	Household size	-
$marital_i$	Marital status	
Se_i	Vector of socio-economic variables	
inc_i	Household income	+
$occup_i$	Occupation	
De_i	Vector of demographic variables	
age_i	Age	U
sex_i	Sex	+(F)
ra_i	Ethnic	
E_i	Vector of variables from the environment	
geo_i	CMSA: codification Bureau of the Census consolidated metropolitan statistical areas	
hab_i	Type of habitat (by size)	+

incorporate contextual effects. To test empirically the predictions of those models, we select from the survey the variables that could potentially reflect those conditions and the outcome achieved. These vectors of explanatory variables are detailed in Table 1.

The number of times that an individual claims to have attended live theatre performances during the last year, y_i , is the dependent variable.⁵

⁵ The reader will find in Table 1 containing a description of the dependent variable and the explanatory variables. We have a count variable that reports on observable market behavior that takes values between 0 and 50.

Explanatory variables related to the stock of cultural capital (S_i) can be classified in different categories (Smith unpublished): general human capital (formed by formal and informal education and skills) and specific human capital for cultural appreciation. First, we may assume that the initial cultural capital endowment could be introduced by the parents' level of education. It is also potentially determined by the public supply of cultural goods and other contextual factors. Since there is no information on this factor, we seek to capture its influence by means of location variables, as explained below. Second, in each period of his/her life, the decision maker has the opportunity to consume cultural goods that will be integrated into his/her stock. To consider the quantities of past cultural consumption transformed into cultural capital, we focus on education level to capture the effect of general human capital and past investments through classes and acting activities.

Apart from consuming the cultural good "theatre performance," those who want to satisfy their cultural need by consuming theatre have other ways of participating (O_i), material goods in the market that differ in some set of attributes (more private and less social, though more flexible). For instance, watching and listening to theater productions on videotape, DVD, TV or radio, or reading dramatic texts are alternative forms of consumption.

As a possible approximation to the distinct precedence of zero variables of the dependent variable we analyze the variable (P_i), which considers the answer to the question: "would you go more to the theatre if you had to face no restrictions?" Borgonovi (2004) uses this variable to estimate satisfaction with the participation function. We do not interpret it in terms of satisfaction, but in terms of constrained or unconstrained maximization and thus in terms of stated preferences as opposed to revealed preferences. Although we initially suspected that there might be an upward bias (people like to please the interviewer), descriptive analysis shows that this attitudinal variable correlates quite well with stated behavior.

As Upright (2004) points out, participation in the arts is a deeply social phenomenon, and therefore not only determined by individual experiences but also by social relationships. Some of the intrinsic intangible values of cultural goods are generated by the consumption of common cultural goods by different members of the community (2001). To capture the influence of the closest background (Sc_i), we introduce information on frequentation by other members of the household. This information is recoded into a dummy variable on the attendance of any other household member.

Alternative uses of personal resources (basically time and money) are work in the marketplace and housekeeping. To accommodate the effect of available time, we recode variables on occupation. To analyze the effect of time devoted to work outside the marketplace, we take household size and marital status into account. Since the decision on the optimal demand for cultural good depends on income, we introduce reported household income. A bundle of socio-economic and demographic characteristics both from the individual and his/her household also enters into the analysis (H_i , Se_i).

To control for further heterogeneity, demographic variables such as age, sex, and ethnic origin are included in the empirical specification (De_i).

Lastly, we have to take into account the influence of contextual factors. The influence of the social dimension of cultural demand is also determined on geographical grounds. Moreover, there could be heterogeneity in the quantity and type of supply (NEA 2004). To include this geographical information (E_i), we use the identifier for the census area (consolidated metropolitan statistical area) and the type of habitat.

The variables included are summarized in Table 1. Table 2 contains descriptive statistics.⁶

3.3 Hypotheses and regularities

Our hypotheses are as follows.

3.3.1 On cultural capital (S_i)

Models that build on the relevance of cultural capital define the stock of available consumption capital as a function of an exogenously determined initial endowment (received in childhood) and previous levels of cultural goods consumption. The higher the education level of parents, the bigger the initial endowment that an agent receives to be integrated into the available stock of consumption capital at a given point, so the higher participation of those agents is expected to be.

In the theoretical framework defined by Stigler and Becker (1977), the available stock of cultural capital is a function of an initial endowment of cultural capital and the vector of past cultural goods consumption. This accumulation function is parameterized by general human capital variables—such as formal education—that would potentially influence the ability to transform resources into commodities. Differences in ability could be captured by education level, allowing for different abilities to interpret past consumption. We have thus accommodated the theoretical reason by which the higher the level of education, the higher the demand, since there will be a bigger available stock of cultural capital (the depreciation rate at which accumulated cultural capital disappears would be lower).

Regarding the accumulation of a specific type of cultural capital for arts appreciation, we compare a bundle of alternative hypotheses. We expect those individuals who receive a particular education in the arts to consume more cultural goods. We can also compare a “duration hypothesis” with an “interest hypothesis.” The former implies that individuals who start to accumulate cultural capital in childhood have more cultural capital available and thus consume more cultural goods; the latter implies that individuals who decide voluntarily to start accumulating specific cultural capital in adulthood consume more cultural goods than those who start in childhood.

⁶ Categorical variables were recoded into dummy variables. Descriptive statistics and estimation results only take into account analytical weights, i.e. the inverse of the probability of being selected. Clusters and strata are not available for confidentiality reasons. For more information on geographical variables and available information, see the technical document from the U.S. Bureau of the Census (2003).

Table 2 Descriptive statistics

Variable	Percentage (mean)	St. deviation
Go (binary)	0.1233	0.0029
Number (count)	0.285	0.0110
Age	45.1270	0.1544
Male	0.4792	0.0045
White	0.8308	0.0036
Black	0.1196	0.0032
Indian	0.0087	0.0008
Island	0.0409	0.0018
inc1 ^a	0.2419	0.0038
inc2	0.1769	0.0034
inc3	0.2543	0.0038
inc4	0.2223	0.0038
inc99	0.1044	0.0027
Full-time worker	0.5454	0.0044
Part-time worker	0.1033	0.0027
Unemployed	0.0371	0.0017
Retired	0.2125	0.0035
Non-active	0.1017	0.0028
Household size	2.8791	0.01376
Married/partner	0.5565	0.01376
Widow	0.0705	0.0021
Single	0.2413	0.004
Divorced	0.1316	0.003
Family go	0.087445	0.0025
Central habitat	0.2338	0.0039
Intermediate	0.4353	0.0044
Non-metropolitan	0.1849	0.0033
Other habitat	0.1463	0.0031
edu1	0.1542	0.0032
edu2	0.3099	0.0041
edu3	0.2762	0.004
edu4	0.1752	0.0034
edu5	0.0844	0.0024
father_edu1	0.4618	0.0045
father_edu2	0.2892	0.004
father_edu3	0.0889	0.0026
father_edu4	0.0991	0.0027
father_edu5	0.0610	0.0021
father_edu99	0.2031	0.0036
mother_edu1	0.418	0.0044
mother_edu2	0.3543	0.0042
mother_edu3	0.1061	0.0028

Table 2 continued

	Variable	Percentage (mean)	St. deviation
	mother_edu4	0.0877	0.0026
	mother_edu5	0.0338	0.0016
	mother_edu99	0.1746	0.0034
	Read theatre	0.0359	0.0017
	Media passive	0.0958	0.0026
	Media active	0.0023	0.0013
	Perform	0.0141	0.0011
	Go more	0.2715	0.0039
	No class	0.9302	0.0023
	class18	0.047	0.0019
	class19	0.031	0.0015

^a The recoded variables that represent the belonging to any of the 4 quartiles of the income distribution were created from a variable that used an interval scale. Thus, as can be seen on the table above, the frequency of each level of the complete dummy variable does not fit exactly with the quartiles

3.3.2 On the variable of going more if possible (P_i)

The variable which informs whether the subject would go to more theater performances if he/she faced no restrictions seems to be a good elicitation of having reached a corner solution at the revealed choice. It informs on whether the agent has reached a constrained or an unconstrained optimum in his/her zero decision and allows us to compare whether those individuals with a higher taste for cultural experience participate more.

3.3.3 On the participation via the media and alternatives to live theater (O_i)

Previous analysis of the SPPA1997 (McCarthy et al. 2001b) reports that there are more people participating in the arts through the media than attending live performances, and that the latter group is bigger than the one engaged in some kind of artistic activity. A big difference between Borgonovi's (2004) work and ours lies in the fact that she uses data on participation in different types of cultural activities, thus taking into account the potential substitutive or complementary effects of one on the others. We concentrate only on those activities that are related to theater and analyze the effect on frequentation of participation through the media and alternatives to live performances in terms of specific education and close goods.

A priori, we can only establish alternative hypotheses. Since our empirical model only takes into account the existence of one cultural good, we cannot infer at a theoretical level what relation exists between different types of participation in the arts. Regarding acting, it could be that people who act have a higher level of very specialized consumption capital, so we might expect them to consume more units of cultural goods; it could also be that people who act produce and consume substitutes for live performances, so they might demand fewer units as consumers. The same ambiguous effect could be proposed for other supports: either those individuals who participate by consuming other theater goods (same good with different attributes)

tend to go less to live performances due to the substitutive nature of these production factors or individuals who consume theater in other forms tend to go more to theater, due to complementarity.

3.3.4 On household income (H_i)

To take household income into account (cultural goods demand is increasing with income, as theoretical models predict that they are normal goods), we create four dummy variables to indicate which quartile of the income distribution household income belongs to under the hypothesis that agents with higher income participate more often.

There is another crucial decision variable in Stigler and Becker's conceptual framework: time. In fact, the optimal allocation of resources problem can collapse into an optimal allocation of time between competing activities: work in the marketplace, housekeeping work, and leisure time (Becker 1965; Gronau 1977; Gronau and Hamermesh 2003). Time devoted to work in the marketplace determines the total amount of available income, time devoted to housekeeping determines what commodities are produced or purchased by an individual, and leisure time is related to the leisure activities that agents can consume and produce. In this broad framework, time is a scant resource that has to be devoted to the production and consumption of commodities to satisfy basic needs, as well as to working hours in the marketplace.

3.3.5 On time devoted to work in the marketplace (Se_i)

We use information from the survey to create five dummy variables to represent occupation. These variables can be ordered in terms of the number of hours spent on work in the marketplace, so we can establish the hypothesis that those individuals who spend less time in the marketplace have more leisure time. However, we should be cautious when deriving total effects: if income effects are cancelled by substitution effects, then those occupations that entail less time working in the marketplace allow for a higher demand for cultural goods. Additionally, we can establish hypotheses on the different effects of the number of hours available for different groups, such as unemployed, retired, and non-active people. According to this dimension of occupation (which does not only take into account the number of hours), we can speak in terms of the discretionary or residual nature of leisure time.

3.3.6 On time devoted to household maintenance (H_i)

The effect of family burdens and ties on the allocation of time operates in the same direction as working time in the marketplace. One of the main differences between leisure time and time devoted to housekeeping is what is called "the first person criterion." For many housekeeping activities the agent can purchase time in the marketplace (in the sense that he/she can substitute hired time for his/her own time

to produce that commodity). However, for the production and consumption of leisure, one's own time is a complement and a necessary factor. In this sense, the relationship between household characteristics and time availability is not that clear. Only if we assume that housekeeping commodities are mainly produced using one's own time can we test the following. Those variables related positively to household needs tend to cause a lower consumption of cultural goods.

The effect of household size may not lead to a clear conclusion on time availability: on the one hand, if the person surveyed is the head of the household, more household members impose a more severe constraint; on the other hand, if he/she is not, the presence of others induces higher per-capita time availability. Marital status may also have a twofold effect. While the presence of a partner may favor the consumption of highly social activities (such as theater going), it could also impose a restriction on the choice set due to the presence of a partner with whom the individual must coordinate.

3.3.7 Other regularities (De_i , E_i)

Traditional control variables for characterizing the socio-economic conditions of individuals are also taken into consideration. Further, we include what Seaman (2005) calls “mixed” factors (gender, race, and ethnicity) that incorporate a measure for socialization variables.

Gray (2003) reports some results and regularities from previous SPPAs. His empirical specification is a model *à la Becker* in which consumption of cultural goods depends on related artistic training and other relevant forms of capital investment in the same spirit in which we derived our first bundle of hypotheses. The direction of the hypothesis for these variables is the same one that we postulate above. However, he also states some hypotheses in terms of out-of-the-model arguments and tests them using SPPA1997 data by estimating binary models.

The first such variable that enters into the analysis is the impact of the place of residence on participation. Gray indicates a positive effect of habitat size due to the higher availability of supply. The foundation of this argument is twofold: the possibility of economies of agglomeration, and the possibility of threshold effects on market size. In observed choice, we can see the effect of consumers who face a lower price of attendance due to lower transportation costs.⁷ We expect theater participation to be higher for those agents living in bigger habitats, bringing further evidence to support the idea that participation in the arts via attendance at live performances is mainly an urban phenomenon.

Regarding age, Gray states that since taste formation is a process that takes time, he would expect participation to increase over the lifetime of the agent. However, estimation does not support this conjecture, since age has no monotone effect.

Another important feature is to analyze the influence of the sex variable in the determination of arts participation. It is a regularity that women participate more

⁷ As wisely suggested by a referee, the fact that supply is higher in bigger cities implies that the total price of attending to performances is smaller.

than men. Gray and other authors systematically argue that the positive and significant effect on participation of being a woman, even after other variables are controlled for, can be explained in terms of early socialization on cultural activities (more exposure to art, while males are socialized rather through sports). Thus, it is a gender effect that informs that this variable is contaminated by educational effects unobserved in our survey. DiMaggio (2004) points out how few studies there are on the role of gender and the formation of cultural capital and participation. In a comparative study on arts participation in USA for 1982–2002 based on successive SPPAs, DiMaggio and Mukhtar (2004) find that feminization persists for all cultural activities and actually increases between 1997 and 2002, despite changes in labor market participation and in traditional household roles. We expect higher participation for women.

In arguments concerned with ethnic origin, the same “mixed” argument applies: differences persist after education and economic variables are controlled for. The effect of these variables varies from one specific artistic activity to another. These features indicate that ethnic origin also captures early socialization effects for each group.

3.3.8 On social capital (Sc_i)

Cultural participation by attendance at live performances is a social activity. Upright (2004) analyzes the relationship of social capital and spousal influence on cultural participation. She finds evidence that supports the following argument for feminization: participation is mainly determined in couples by women. Those men whose wives are more highly educated and have more cultural capital are more likely to participate when their wives do. We are interested in shaping all the sample behavior, so we proceed in a different way: we seek to study the influence of the interdependency of behavior at an intra-household level trying to determine whether there is a relationship between observed behavior and participation of other household members.

4 Estimation methods

Taking into account the hypotheses defined above and the variables in the dataset, we propose to estimate the following individual participation equation with the variables as defined in Sect. 3.

$$y_i = f(\mathbf{x}_i) = f(S_i, P_i, O_i, Sc_i, H_i, Se_i, De_i, E_i) \quad (2)$$

Our estimation process starts by recognizing the count data nature of our dependent variable.⁸ Throughout the estimation procedure, we provide arguments

⁸ We could also treat this variable as an ordered variable as Boronovi (2004) does. When commenting our final results, we will compare them with the ones that we obtain if we estimate an ordered logit model (assuming that the parallel regression assumption is fulfilled).

that justify both the selection of the model (econometrically) and the suitability of the model for explaining participation phenomena.⁹

The outline goes as follows. We start by modeling our count data, assuming that all data are generated by the same data-generating process. To that end we estimate a Poisson model and, since the equidispersion assumption that has to be met is violated, we estimate a Negative Binomial model. Still, the fit of the model is not good, especially for the zero outcome: there is evidence that zero observations may be driven by two different data-generating processes. This is why we estimate a Zero Inflated Poisson (ZIP) and then a Zero Inflated Negative Binomial Model.

To explain the process of fit of the different estimation models, we include a subsection in the Appendix that presents the goodness of fit tests and measures, following a procedure analogous to the research by Cameron and Trivedi (1998). At each point in time, we explicitly mention the behavioral assumptions of fitting such a model.

To estimate the model in Eq. 2, we estimate a count model. Thus, we start with the simplest one: a Poisson Regression Model. The subsequent goodness of fit test indicates that the equidispersion assumption is violated, so a Negative Binomial Regression would improve matters. This is a common phenomenon when unobserved heterogeneity explains a major part of the dependent variable. The estimated parameter for the overdispersion of variance is significant, confirming that heterogeneity outside our empirical specification is affecting the estimations. However, the fit of the model is not good, providing evidence of additional unobserved heterogeneity.¹⁰ As noted above, this individual heterogeneity induces (i) overdispersion, thus violating the equidispersion assumption and (ii) overrepresentation of zero values.

The assumptions underlying these individual decision making are the following: assume that there are two sub-populations in the population under study:

⁹ At each point in time, we explicitly mention the behavioral assumptions of fitting such a model. We do not consider in depth two other major approaches: hurdle models and latent class (Mullahy 1986). The use of a two-part hurdle models (TPM) would imply that all zeros are ruled by the same process; the use of a latent model (LCM) would be useful to classify types of attendees. TPM and LCM have been widely used to estimate heterogeneous individual behavior in other areas, mostly in health related choices (as diverse as health services use or consumption of unhealthy substances, e.g. tobacco demand) or in demand for recreation. From a statistical point of view, the TPM is also a finite mixture with a degenerate component. It combines zeros from a binomial density with the positives from a zero truncated density. For the implications of following the TPM approach and estimation procedures, see the outstanding contribution by Mullahy (1998). In the appendix, the reader will find the goodness of fit measures of a hurdle model. The LCM is more flexible because it permits mixing with respect to both zeros and positives (Deb and Trivedi 2002). It is not clear a priori which model would perform better empirically (see, for example, Jiménez-Martín et al. 2002).

¹⁰ So that the process of the reshaping of the functional form to be estimated can be followed, goodness of fit tests and measures are provided in the Appendix. Those results justify the choice of the Zero Inflated Negative Binomial presented below as that which best represents individual behavior. We also provide the goodness of fit statistics of a Hurdle model (thus, a two-part model) that consider that there are two processes determining the attendance: a logit model determines the probability of going or not going and a Truncated Negative Binomial to determine the positive count outcomes. Note that this Hurdle model assumes that all the zero outcomes are generated by the same process.

- *Those individuals that never go to the theater and do not even consider the possibility.* This may be, for instance, because they do not like it, or because they have been socialized under other forms of artistic participation (e.g. music) and have thus accumulated a specific stock of cultural capital to satisfy their cultural appreciation needs in a more efficient way.
- *Those who may or may not go.* The behavior of individuals in this sub-population may diverge depending on resource availability: leisure time, financial constraints, skills needed to understand and appreciate the play, etc. We can establish that the observed behavior of an individual belonging to this group who demands zero units of cultural good reveals a corner solution as the constrained optimum.

To analyze the differences in behavior of these two groups using the data derived from the survey, we must be aware that it is not possible to determine whether a zero observation comes from one sub-population or the other. Instead of using an arbitrary criterion, such as modeling the zero outcome as a hurdle, with the conditional distribution of the positive values being governed by a zero-truncated count model, we decided to estimate the zero inflated model using Poisson and Negative Binomial.

For analytical purposes, we define an indicator function A that takes the value 1 if the agent belongs to the group that never goes to the theater and 0 if he/she belongs to the other group. This belonging is a binary event that may be modeled as a logit or probit model. We propose the latter,

$$\psi_i = \Pr(A_i = 1 | z_i) = F(z_i \gamma) \quad (3)$$

$$\psi_i = \frac{\exp(\gamma_1 z_1 + \cdots \gamma_m z_m)}{1 + \exp(\gamma_1 z_1 + \cdots \gamma_m z_m)} \quad (4)$$

where ψ_i is the probability of individual i belonging to group A . Explanatory variables z_i refer to the variables explaining the inflation of zeros. Parameter vector γ captures the influence of explanatory variables on the binary outcome of “always zero” or not. If any of the available variables is the one determining this outcome, then the analysis turns into the standard binary regression. However, since the fact of belonging to a given group is a latent variable, we cannot know if an individual belongs to to one group or the other.

Among those individuals who are not “always zero,” i.e. those for whom there is a positive probability that they may go to a theater, the probability of each of the observed counts (in terms of number of times, including zeros derived from corner solutions) is determined by a Negative Binomial Regression.

These two behavioral models are conditional on variables z and x .

$$\mathbf{z}_i = \mathbf{x}_i = (S_i, P_i, O_i, Sc_i, H_i, Se_i, De_i, E_i) \quad (5)$$

We want to keep the same set of explanatory variables for the binary outcome and for the count outcome. This allows us to compare the effect of each of the explanatory variables in each part of the behavioral proposal.

$$\Pr(y_i|x_i, A_i = 0) = \frac{\Gamma(y_i + \alpha^{-1})}{y_i! \Gamma(\alpha^{-1})} \left(\frac{\alpha^{-1}}{\alpha^{-1} + \mu_i} \right)^{\alpha^{-1}} \left(\frac{\mu_i}{\alpha^{-1} + \mu_i} \right)^{y_i} \quad (6)$$

$$\mu_i = \exp(x_i \beta)$$

These two groups are mixed to determine the population model taking into account the proportions of each group, which are defined by

$$\Pr(A_i = 1) = \psi_i, \quad \text{and} \quad \Pr(A_i = 0) = 1 - \psi_i$$

while the probability of an event being zero for each group is defined by

$$\Pr(y_i = 0|A_i = 1, x_i, z_i) = 1 \text{ by the definition of group A} \quad (7)$$

$$\Pr(y_i = 0|A_i = 0, x_i, z_i) = \text{the result of the Negative Binomial Regression}$$

Combining previous statements, we can obtain the total probability of an event being zero (which adds up zeros coming from two distinct decision models)

$$\Pr(y_i = 0|x_i, z_i) = \{\psi_i x 1\} + [1 - \psi_i] x \Pr(y_i = 0|x_i, A_i = 0) \quad (8)$$

$$= \psi_i + \{(1 - \psi_i) x \Pr(y_i = 0|x_i, A_i = 0)\}$$

Any event different from zero can only belong to an individual on the group with positive probability of attendance.

$$\Pr(y_i = k|x_i, z_i) = \{\psi_i x 0\} + [1 - \psi_i] x \Pr(y_i = k|x_i, A_i = 0) \quad (9)$$

$$(1 - \psi_i) x \Pr(y_i = k|x_i, A_i = 0)\}$$

Each expect count is determined by

$$E(y|x, z) = [0 x \psi] + \{\mu x (1 - \psi)\} = \mu(1 - \psi)$$

Since $0 < \psi < 1$, the expected value of any count will be lower than μ (conditional mean), showing that the structure of means for zero inflated models is different from Poisson or Negative Binomial models. Therefore, we shall estimate a *Zero Inflated Negative Binomial*, since this will allow us to control for the double effect of individual unobserved and omitted heterogeneity.

To control for the contextual influences of the environment in which the individual takes her decisions, we will estimate robust variance-covariance matrix by using clusters defined by CMSA.

5 Estimation results

We have characterized two distinct types of behavior in the market.

- A group that never goes to the theater: belonging to this group is characterized by a binary outcome
- Another group that may demand theater performances: if an agent from this group reveals a zero frequentation in the market, this is caused by some restriction in his/her constrained maximization problem that induces a corner solution. We characterize what increases the probability of going more times.

We start by discussing the determinants that reduce the probability of never attending, then discuss the influence of the variables on the frequency of participation and end up by relating the two results with previous findings and with the results derived from estimating the same specification using an ordered logit model.¹¹

Men are more likely never to participate in theater performances. Income has a monotone effect in reducing this probability of never attending for the first three quartiles of the income distribution. Being single reduces this probability. Own formal education has the expected effect but only for the very basic level, and we do not find any statistically significant effect of parental education except for very high levels in the part of the father (we suspect that the influence of this variable on family wealth could be leading this result). All the variables related to other kinds of theater consumption or participation have a greater impact in reducing the probability of never attending a theater performance.

For those people who have a positive probability of attending a theater performance, we find that age increases the probability of going more; the same applies to sex (females go more). The effect of income is only statistically significant for the last quartile of the distribution. Once again, being single leads to going more times and living in an urban area also has a positive effect. Low education levels determine low frequency and the highest levels of education make a difference. Reading theater, consuming it on the media (in a passive way) and having received drama classes as an adult also lead to higher frequentation. If a person in this subgroup reported that he/she would like to go more, then that person is more likely to participate more frequently.

To clarify these results, we now examine jointly the influence of the proposed variables for the behavior of the two subgroups and put them in perspective by using the most closely related results, those reported by Borgonovi (2004).¹²

Regarding age, we have reached a different result. Borgonovi concludes that theater is a particular art form that appeals to younger generations, a feature specific to this form of arts participation. However, we have found a different age profile. We find that age has a positive effect on both participation and frequency, and we find no evidence in support of the exceptional nature of this type of artistic activity. If instead of taking into account the count data nature of our dependent variable we take into account its ordinality, the results derived from an ordered logit model would match her previous result.

We find evidence that supports the feminization hypothesis for the two types of behavior, which is contrary to the finding of Borgonovi, who concludes that females tend to be more likely than males to attend, but they do so as occasional visitors.

¹¹ Caveat: as mentioned above, the parallel assumption is not met, so we cannot trust these estimations. Rather, a generalized ordered model would perform better, in the spirit of what Borgonovi herself estimates. However, we would then encounter a new problem: we would seldom be able to estimate such a model considering so many outcomes of the dependent variable. In our sample, there are a great many values of the dependent variable with very few observations. One possible way to overcome this problem is to regroup under an arbitrary criterion. Another way is to estimate a latent class model.

¹² We would like to remark that very few variables are statistically significant in Borgonovi's specification for "drama" participations in Table 5 and 6. Sometimes the statistically significant coefficients are small in size (odds ratio around 1).

Other results indicate evidence that habitat size determines frequency of participation but has no statistically significant effect on the probability of never going. Being single seems to determine that the probability of never going is reduced and the frequency of attendance increases. In a household decision framework, it seems that this result sheds some light on the discretionary nature of leisure time for these agents. When estimating the effect of this same set of explanatory variables using an ordered logit, the results are somewhat different: the bigger the household, the lower the probability of participating more; if one member of a household participates, then it is much more likely that others will participate more (our estimated odds ratio is around 5). The effect of family participation disappears when estimating the ZINB (only some evidence of very marginal significance for the count model) (Table 3). These results suggest that family participation is probably not so crucial in explaining adults' attendance.¹³

Our results for educational variables help us to analyze the formation of cultural capital and its relevance. One could distinguish between specific cultural capital and general cultural capital acquired by education (recall that taste formation models predict that more educated individuals can be more productive when accumulating their cultural capital). Reports in the USA by the National Endowment for the Arts (2004), establish the significance of educational factors with respect to economic ones in seeking to explain participation. We can support the importance of the role of education: in the binary part of the model, education is a factor that increases participation, since the probability of never going decreases as education increases (this is a monotone result, except for the lowest level); in the count model, there is also a positive relationship between education level and demanded quantity.

Participation in art education is much more important than any other type of general education. This is a result from Borgonovi's study and from ours. The complementary effect in our case rules for "active" participation in drama (acting) and for participation by consuming closely related goods. We also find that the coefficients for such variables have the highest absolute value and that the effect is stronger when we want to characterize the behavior of the never goers. Thus, this seems to provide further evidence that many forms of participation in the performing arts are complements. Recall that we presented three alternative definitions of participation—attendance, consumption via the media, and practice of activities. It seems that participation in one particular art form increases the likelihood of participation in others and that participation in one of the three ways detailed above increases the likelihood of attendance.

Family background (in terms of mother's and father's education level, incorporated as a measure of the initial endowment that a child receives under the cultural capital hypothesis) does not seem to play a major independent effect on participation in performing arts. This is the same result that Borgonovi reaches.

There are other variables that try to examine the effect of available personal resources on never going and on frequency of participation. Income seems to have a more important effect on determining the likelihood of the access to participation. If we treat our dependent variable as an ordinal outcome, we find a monotone

¹³ With the exception of those studies that determine spouses' joint participation.

Table 3 ZINB estimation results^a

Variable	Always zero			Count model		
	$\hat{\gamma}$	<i>P</i> -value	$e^{\hat{\gamma}}$	$\hat{\beta}$	<i>P</i> -value	$e^{\hat{\beta}}$
Age	−0.00150	0.374	0.9985	−0.08663**	0.000	0.9170
Age ²	0.00001	0.248	1.0000	0.00198**	0.000	1.0020
Age ³	0.45013**	0.000	1.5685	−0.00001**	0.004	1.0000
Male	0.67583**	0.004	1.9657	−0.14172*	0.087	0.8679
Black	−0.99418	0.234	0.3700	0.28428	0.253	1.3288
Indian	0.07422	0.862	1.0770	−0.65652*	0.083	0.5187
Island	−0.38723*	0.087	0.6789	−0.05386	0.869	0.9476
inc2	−0.56736**	0.010	0.5670	0.03375	0.859	1.0343
inc3	−0.78865**	0.000	0.4545	0.17934	0.250	1.1964
inc4	−0.40183*	0.073	0.6691	0.31410*	0.054	1.3690
inc99	−0.06239	0.760	0.9395	0.26764	0.120	1.3069
Part-time	−0.52630	0.192	0.5908	0.06728	0.561	1.0696
Unemployed	0.37137	0.109	1.4497	−0.34491*	0.074	0.7083
Retired	0.11367	0.641	1.1204	0.10779	0.473	1.1138
Non-active	−0.57637	0.114	0.5619	0.08407	0.549	1.0877
Household size	−0.58172*	0.091	0.5589	−0.02849	0.480	0.9719
Widow	−0.27928	0.310	0.7563	−0.08681	0.633	0.9168
Single	−0.67334**	0.006	0.5100	0.36778**	0.019	1.4445
Divorced	−0.26631	0.215	0.7662	−0.06679	0.649	0.9354
Family go	0.20292	0.244	1.2250	0.17363	0.104	1.1896
Central	−0.11116	0.349	0.8948	0.455555**	0.001	1.5770
Intermediate	0.04348	0.771	1.0444	0.11448	0.180	1.1213
Other habitat	2.25565**	0.033	9.5415	0.09282	0.396	1.0973
edu1	−0.59669**	0.001	0.5506	−0.99695**	0.003	0.3690
edu3	−0.63959**	0.001	0.5275	−0.03860	0.793	0.9621
edu4	−0.92931**	0.000	0.3948	0.25538*	0.089	1.2910
edu5	0.11323	0.115	1.1199	0.50686**	0.001	1.6601
father_edu1	−0.13340	0.649	0.8751	−0.19304	0.282	0.8244
father_edu3	−0.44618	0.123	0.6401	−0.15042	0.359	0.8603
father_edu4	−0.80822**	0.011	0.4457	−0.12420	0.433	0.8832
father_edu5	0.91225**	0.001	2.4899	−0.11028	0.480	0.8956
father_edu99	0.33916	0.111	1.4038	0.77984**	0.001	2.1811
mother_edu1	−0.09049	0.674	0.9135	0.32326*	0.052	1.3816
mother_edu3	−0.32305	0.184	0.7239	0.0392	0.749	1.0401
mother_edu4	−0.28790	0.186	0.7498	0.01469	0.914	1.0148
mother_edu5	−0.42490	0.361	0.6538	0.10195	0.421	1.1073
mother_edu99	−0.51399	0.127	0.5981	−0.67554	0.005	0.5089
Theatre read	−1.01301**	0.000	0.3631	0.75561**	0.000	2.1289
Media passive	−1.42370**	0.008	0.2408	0.23444**	0.026	1.2642
Media active	−2.67922**	0.018	0.0686	−0.00042	0.997	0.9996
Perform	−1.15885**	0.000	0.3138	0.22223	0.219	1.2489
class18	−0.46542	0.211	0.6279	0.05960	0.626	1.0614
class19	−3.16351**	0.000	0.0423	0.44146**	0.000	1.5550
Go more	−0.38540	0.181	0.3602	0.17363**	0.001	1.1896

Table 3 continued

Sample size	$N = 16659$	lnalpha	0.09217
Log-likelihood	$-9.032e+11$	Alpha	1.09656
Pseudo- R^2	0.198	SE (alpha)	0.11723

^a Significance level: *10%; **5%

Note: $\hat{\gamma}$: estimated coefficient; $e^{\hat{\gamma}}$: factor change in probability (odds) for a unit increase of x ; $\hat{\beta}$: estimated coefficient; $e^{\hat{\beta}}$: factor change in expected count for a unit increase of x

Omitted categories: female, white, inc1, full-time, married, family doesn't go, no metropolitan, edu2, father-edu2, mother_edu2, noperform, wouldn't go more, no classes

increasing effect of income. Borgonovi reports a positive effect for participation and also a positive, albeit very small, effect for frequency, (odds ratio = 1.067).

Finally, we comment on the estimations for the variable that captures whether the respondent would attend more if he/she had no constraints. There is a barely statistically significant effect for the count part of the model. However, when estimating the ordered logit, the estimated odds ratio is 2.86, one of the highest effects.

6 Conclusions

One of the most consistent findings in the literature on arts participation has been the positive relationship between participation and income and education and the relatively small fraction of population that participates (count participation variables are contaminated by the “excess zeros” problem). Why is it that so many people determine that the optimal choice is not to reveal any preference for these cultural goods? If we inferred that all no-goers are simply not interested in the arts, we might be making a big mistake. The main contribution of this piece of research is the characterization of the behavior of two heterogeneous populations: those who never attend a theater performance and those who may go a positive number of times. These two groups are endogenously determined by the model itself. Our estimation method takes into account the behavioral assumptions derived from the individual decision-making process of maximization of utility subject to some constraints. We estimate the frequentation model for theater performances in US by means of a ZINB. Unobserved heterogeneity not included in the proposed empirical specification is inducing two points in the data: on the one hand overdispersion with respect to the Poisson Regression (the value of the conditional variance exceeds the value of the conditional mean), and on the other hand over-representation of zero values, since they come from two distinct populations. Therefore, the proposed model has the advantage of latent classification in two different behavioral models that are jointly estimated by the mixture of two data generation processes.

One of the results traditionally reported in the literature on arts participation is the greater significance of cultural variables than economic ones in explaining different levels of attendance. In our study we find that economic variables work as

deterrents to participation, whereas variables measuring different aspects of cultural capital explain both access and intensity of participation. We provide evidence on the role of individual cultural capital in explaining culture consumption. It turns out that own general cultural capital (accumulated by means of formal education) influences never going, but is less influential on the amount of times. Additionally, it seems that it is specific cultural capital rather than generic education that has the greater influence. We also find supportive evidence for the complementary nature of different ways of participation.

Lastly, as in previous studies, we find additional evidence for the characterization of participation in theater as an urban, mainly feminine phenomenon; we believe that these results can be explained in terms of specialization in the consumption of certain cultural goods. Potentially, much of the unobserved heterogeneity due to informal socialization patterns transmitted in the household might be captured by these variables. The existing literature has explained them by extra-economic arguments (such as early socialization in the arts for women, for instance), but in our framework this result may imply the existence of more productive agents. It could be the case that women have a higher external initial capital endowment, and thus they may be characterized by technology leading to higher consumption.

There is still a lot of work to be done in an attempt to find better ways of estimating participation functions that may shed some light on the puzzling questions of the determinants of attendance and of non-participation.

Acknowledgments I would like to thank the guidance of Antonio Morales-Siles, who was my thesis advisor, as well as helpful comments by the Editor and two anonymous referees. This paper also benefited from comments received on the 14th Cultural Economics International Conference, especially from the discussion of Bruce Seaman. All errors are solely mine. I would also like to thank the financial support of SGAE-Fundacion Autor, the Spanish Ministry of Education (SEJ 2005/06099) and the Basque Government (BFI-05.225 and IT-241-07).

Appendix

Model selection criteria

At every stage of the estimation of the count model, we have performed several contrasts and goodness of fit tests to select the most suitable model. We show the path followed to select the ZINB Model as the sample data-generating process (Long and Freese 2006).

The less sophisticated model is the *Poisson Regression Model*; it has to fulfill the condition of equal conditional mean and variance (*equidispersion*). Since the goodness of fit test rejects the hypothesis of equidispersion (the conditional variance would be higher than the conditional mean), we estimate a *Negative Binomial II Model* (Cameron and Trivedi 1998),¹⁴ which parameterize the conditional variance so it can accommodate the overdispersion. Thus, we estimate a zero inflated model

¹⁴ In what follows, when talking about the Negative Binomial Model, we will refer to the Negative Binomial II Model, in terms of Cameron and Trivedi (1998).

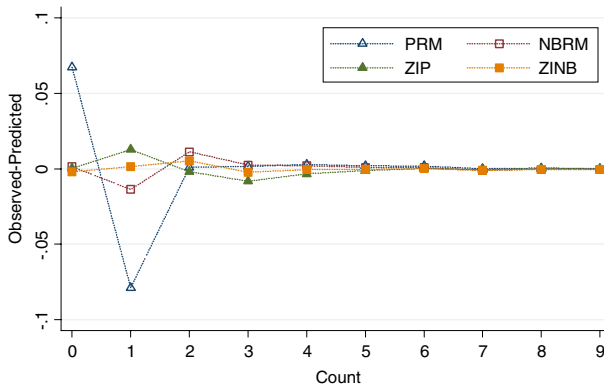


Fig. 1 Difference between observed and predicted values (truncated at 9 for this graphic). *Note:* Positive deviations show underpredictions

that would accommodate simultaneously the binary fact of participation as well as the count model. For the binary participation decision, we estimate a logit regression model, whereas for the count we face a double alternative: Poisson or negative binomial distribution. In this way, we have two models, the ZIP and the ZINB, being the second model a particular case of the first one. We also estimate a *Hurdle* model (logit for the binary outcome and *Negative Binomial II* for the count).

Goodness of fit statistics for the estimated models

	BIC	AIC	Loglik
Poisson	2.453e+12	1.472e+08	-1.226e+12
NB (Negative Binomial II)	1.908e+12	1.146e+08	-9.542e+11
ZIP	1.950e+12	1.171e+08	-9.752e+11
ZINB	1.806e+12	1.084e+08	-9.032e+11
Hurdle	NA	1.08e+08	-9.036e+11

Goodness of fit tests for the estimated models^a

	Statistic	Test
Poisson	$\chi^2 = 1.84e+12$	$P > \chi^2 (16,656) = 0.0000$
Poisson versus NB	$LR = -2 * (llp - llnb) = 5.4e+11$	with $P \geq 0.000$
ZIP versus ZINB	$LR = -2 * (llzip - llzinb) = 1.4e+11$	with $P \geq 0.000$

^a We could not estimate Vuong tests to discard Poisson versus ZIP and NB versus ZINB since we used probability weights to estimate our models. Vuong test cannot be computed in stata with pweight option

As seen in the tables above, standard information criteria provide strong evidence to support the suitability of the *ZINB*. Finally, we present the graphic that represents how well each of the models have performed (Fig. 1).

References

- Abbé-Decarroux, F., & Grin, F. (1992). Risk, risk aversion and the demand for performing arts. In R. Towse & A. Khakee (Eds.), *Cultural economics* (pp. 125–140). Springer-Verlag.
- Ateca-Amestoy, V. (2007). Cultural capital and demand. *Economics Bulletin*, 26(1), 1–9.
- Becker, G. S. (1965). A theory of the allocation of time. *Economic Journal*, 75(September), 493–517.
- Borgonovi, F. (2004). Performing arts: An economic approach. *Applied Economics*, 36, 1871–1885.
- Bourdieu, P. (1988). *La Distinción. Criterios y Bases Sociales del Gusto*. Spanish edition: (trans: Ruiz de Elvira, M.C. Taurus).
- Bureau of the Census. (2003). *Current Population Survey, August 2002: Public Participation in the Arts Supplement. Technical Documentation Attachments*. Realized by the Bureau of the Census for the National Endowment for the Arts, Washington.
- Cameron, C. A., & Trivedi, P. K. (1998). *Regression analysis of count data*. Cambridge University Press.
- Deb, P., & Trivedi, P. K. (2002). The structure of demand for health care latent class versus two part models. *Journal of Health Economics*, 21(4), 601–625.
- DiMaggio, P. (2004). Gender, networks, and cultural capital. *Poetics*, 32, 99–103.
- DiMaggio, P., & Mukhtar, T. (2004). Arts participation as cultural capital in the United States, 1982–2002: Signs of decline? *Poetics*, 32, 169–194.
- DiMaggio, P. J., & Ostrower, F. (1990). Participation in the arts by Black and White Americans. In D. Pankratz & V. Morrie (Eds.), *The future of the arts: Public policy and arts research* (pp. 105–140). Praeger Press.
- Fernández-Blanco, V., Prieto-Rodríguez, J., & Orea-Sánchez, L. (2004). *Movie enthusiasts versus cinemagoers in Spain: A latent class model approach*. Presented at 13th ACEI Conference. University of Illinois at Chicago.
- Frey, B. S. (2003). *Arts & economics: Analysis & cultural policy*. New York: Springer Verlag.
- Gapinski, J. H. (1986). The lively arts as substitutes for the lively arts. *American Economic Review*, 76(2), 20–25.
- Gray, C. M. (1995). Turning on and tuning. In Media participation in the arts. *National Endowment for the Arts Research Division Report #33*. Seven Locks Press.
- Gray, C. M. (2003). Participation. In R. Towse (Ed.), *A handbook of cultural economics*. Edward Elgar.
- Gronau, R. (1977). Leisure, home production and work—the theory of the allocation of time revisited. *Journal of Political Economy*, 85, 1099–1123.
- Gronau, R., & Hamermesh, D. S. (2003). Time vs. goods: The value of measuring household production technologies. *NBER Working Paper 9650*. National Bureau of Economic Research.
- Jiménez-Martín, S., Labeaga, J. M., & Martínez-Granado, M. (2002). Latent class versus two-part models in the demand for physician services across the European Union. *Health Economics*, 11, 301–321.
- Lévy-Garboua, L., & Montmarquette, C. (1996). A microeconomic study of theater demand. *Journal of Cultural Economics*, 20, 25–50.
- Long, J. S., & Freese, J. (2006). *Regression models for categorical dependent variables using stata* (2nd ed.). Stata Press.
- McCarthy, K. F., Brooks, A., Lowell, J., & Zakaras, J. (2001a). *The performing arts in a new era*. RAND Corporation.
- McCarthy, K. F., Ondaatje, E. H., & Zakaras, L. (2001b). *Guide to the literature on participation in the Arts*. RAND Corporation.
- Moore, T. G. (1966). The demand for Broadway theater tickets. *Review of Economics and Statistics*, 48(1), 79–97.
- Morrison, W. G., & West, E. G. (1986). Child exposure to the performing arts: The implication for adult demand. *Journal of Cultural Economics*, 10(1), 17–24.
- Mullahy, J. (1986). Specification and testing of some modified count data models. *Journal of Econometrics*, 33, 341–365.
- Mullahy, J. (1998). Much ado about two: Reconsidering retransformations and the two-part model in health econometrics. *Journal of Health Economics*, 17, 247–281.
- National Endowment for the Arts. (2004). *2002 Survey of Public Participation in the Arts. Research Division Report #45*. National Endowment for the Arts.
- Sable, K., & Kling, R. (2001). The double public good: A conceptual framework for “Shared Experience” values associated with heritage conservation. *Journal of Cultural Economics*, 25, 77–89.

- Seaman, B. A. (2005). *Attendance and public participation in the performing arts: A review of the empirical literature*. Andrew Young School of Policy Studies Research Paper Series No. 06–25. Georgia State University.
- Smith, T. M. The demand for culture; A human capital approach. In *Two essays on the economics of the arts: The demand for culture and the occupational mobility of artists*. Unpublished Doctoral Thesis, University of Illinois at Chicago.
- Stigler, G., & Becker, G. S. (1977). De Gustibus Non est Disputandum. *Journal of Political Economy*, 67(1), 76–90.
- Upright, C. B. (2004). Social capital and spousal participation: Spousal influences on attendance at arts events. *Poetics*, 32, 129–143.
- Veblen, T. (1965) *The theory of the leisure class*. Augustus M. Kelley, (1965 reedition).